Resolution and Noise Properties of MAP Reconstruction for Fully 3D PET

[†‡]Jinyi Qi, *Member, IEEE* and [‡]Richard M. Leahy, *Member, IEEE*
[†]Center for Functional Imaging, Lawrence Berkeley National Laboratory,
University of California, Berkeley, CA 94720
[‡]Signal and Image Processing Institute, University of Southern California,
Los Angeles CA 90089-2564

Corresponding author:
Richard M. Leahy
Signal and Image Processing Institute
3740 McClintock Ave, EEB400
University of Southern California
Los Angeles CA 90089-2564
Tel: (213) 740-4659
Fax: (213) 740 4651
e-mail: leahy@sipi.usc.edu.

This work was supported by the National Cancer Institute under Grant No. R01 CA59794.

Abstract

We derive approximate analytical expressions for the local impulse response and covariance of images reconstructed from fully-3D PET data using MAP (maximum *a posteriori*) estimation. These expressions explicitly account for the spatially variant detector response and sensitivity of a 3D tomograph. The resulting spatially variant impulse response and covariance are computed using 3D Fourier transforms. A truncated Gaussian distribution is used to account for the effect on the variance of the non-negativity constraint used in MAP reconstruction. Using Monte Carlo simulations and phantom data from the microPET small animal scanner, we show that the approximations provide reasonably accurate estimates of contrast recovery and covariance of MAP reconstruction for priors with quadratic energy functions. We also describe how these analytical results can be used to achieve near uniform contrast recovery throughout the reconstructed volume.

Keywords

MAP estimation, image reconstruction, resolution analysis, covariance, uniform resolution, positron emission tomography, fully 3D PET.

I. INTRODUCTION

MAP image reconstruction methods can combine accurate physical models for coincidence detection in 3D PET tomographs and statistical models for the photon-limited nature of the coincidence data with regularizing smoothing priors on the image. As we have previously shown [1], [2], this translates into improved resolution and noise performance when compared to filtered-backprojection (FBP) methods that are based on a simpler line-integral model and do not explicitly model the noise distribution.

Fessler and Rogers [3] have shown that MAP (or equivalently, penalized maximum likelihood) reconstruction produces images with object-dependent resolution and variance for 2D PET systems with a spatially-invariant response. The situation is further complicated when the true spatially variant sinogram response is considered [1]. In 3D PET systems the large axial variation in sensitivity produces increased spatially variant behavior. The utility of the MAP approach for 3D PET would be enhanced if we were able to characterize this spatially variant behavior through computation of the resolution and covariance of the resulting images. These computations should include the effects of both axial variation in sensitivity and spatially variant sinogram response.

Here we develop approximate analytical expressions for the local impulse response and covariance of 3D MAP images. These results can be used not only to characterize the images, but also to modify the smoothing effect of the prior to optimize performance for specific tasks. For instance, in combination with computer observer models, these results have been used to compute ROC curves for lesion detectability, and in turn, to optimize MAP reconstruction for lesion detection [4]. Here we show an example of using our local impulse response analysis to develop a method to spatially adapt the smoothing prior, as proposed for the 2D case in [3], to achieve near uniform contrast recovery throughout the scanner field of view.

Since most iterative algorithms for PET, and in particular the MAP method in [1], are nonlinear, the statistical properties of the reconstructions cannot be computed directly from those of the data and approximations are typically required to make the problem tractable. Barrett *et al* [5] and Wang *et al* [6] have derived approximate expressions for the bias and covariance of EM and generalized-EM algorithms as a function of iteration. This approach is very useful for algorithms that are terminated at early iterations but computation cost is high and the accuracy of the approximation can deteriorate at higher iterations.

An alternative approach for algorithms that are iterated to effective convergence is to analyze the properties of the images that represent a fixed point of the algorithm [3], [7], [8]. Building on this work, we have derived simplified theoretical expressions for the local impulse response and the voxel-wise variance of MAP reconstruction for 2D PET systems [9]. The resulting expressions are readily computed using 2D discrete Fourier transforms and their relatively simple algebraic form reveals the effect of the prior smoothing parameter on image resolution and variance. In [10] we extended these results to approximate the full image covariance and also described a method for using a truncated-Gaussian model to account for the effect of the non-negativity constraint on image variance. All of these previous studies [3], [7], [8], [9], [10] were restricted to 2D PET and assumed a shift invariance in the combined forward and back projection operators, which is not applicable to fully 3D PET.

Here we extend the results in [9] and [10] to fully 3D PET. In this work we include the effects of spatially variant sinogram response [1], variations in sensitivity due to "missing" projections, and the non-negativity constraint. Resolution is studied using a local "contrast recovery coefficient" (CRC) computed at each voxel using the local impulse response [3]. Analytic expressions for contrast recovery and covariance reveal the source of spatial variations in these quantities and the effect of the smoothing parameter. Using these simplified expressions we can directly control the resolution vs. noise trade off. For example we can spatially adapt the smoothing parameter to achieve a specific variance or contrast recovery value or, as proposed in [9], maximize contrast to noise ratio to optimize reconstructions for lesion detection. We note that when the smoothing term is made data-adaptive, the algorithm ceases to be a true Bayesian method. However, the spatially variant smoothing weights are computed before the image is reconstructed; the image can then be reconstructed using these weights with the same program that we use to compute true MAP estimates. Although this paper deals with PET image reconstruction, the techniques presented below represent a general approach for analyzing images computed from space-variant systems using MAP estimators.

II. BACKGROUND

A. MAP Reconstruction

PET data are well modeled as a collection of independent Poisson random variables with the log-likelihood function

$$L(\boldsymbol{y}|\boldsymbol{x}) = \sum_i y_i \log \bar{y}_i - \bar{y}_i - \log y_i! \tag{1}$$

where $\boldsymbol{x} \in R^{N\times 1}$ is the unknown image, $\boldsymbol{y} \in R^{M\times 1}$ the measured sinogram, and $\bar{\boldsymbol{y}} \in R^{M\times 1}$ the mean of the sinogram. The mean sinogram $\bar{\boldsymbol{y}}$ is related to the image, $\boldsymbol{x}$, through an affine transform

$$\bar{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{x} + \boldsymbol{s} + \boldsymbol{r} \tag{2}$$

where $\boldsymbol{P} \in I\!R^{M\times N}$ is the detection probability matrix , and $\boldsymbol{s} \in I\!R^{M\times 1}$ and $\boldsymbol{r} \in I\!R^{M\times 1}$ account for the presence of scatter and randoms in the data, respectively.

When operated in standard mode, PET scanners pre-correct for randoms by computing the difference between coincidence events collected using a “prompt” coincidence timing window and those in a delayed timing window of equal duration. This correction method is based on the assumption that the events in the delayed timing window have mean equal to that of the randoms in the prompt timing window. The precorrected data $\boldsymbol{y}$ has mean $\boldsymbol{P}\boldsymbol{x} + \boldsymbol{s}$ and variance $\boldsymbol{P}\boldsymbol{x} + \boldsymbol{s} + 2\boldsymbol{r}$, so a Poisson model does not reflect the true variance. The true distribution has a numerically intractable form, however, the shifted-Poisson model with log likelihood [1]

$$L(\boldsymbol{y}|\boldsymbol{x}) = \sum_{i=1}^{M} (y_i + 2r_i) \log\left((\boldsymbol{P}\boldsymbol{x})_i + s_i + 2r_i\right) - \left((\boldsymbol{P}\boldsymbol{x})_i + s_i + 2r_i\right) \tag{3}$$

serves as a good approximation [11].

The detection probability matrix $\boldsymbol{P}$ can be accurately modeled using the factored detection probability matrix that we developed in [12] and [1]

$$\boldsymbol{P} = \boldsymbol{P}_{det.sens} \boldsymbol{P}_{det.blur} \boldsymbol{P}_{attn} \boldsymbol{P}_{geom} \tag{4}$$

where $\boldsymbol{P}_{geom}$ is the geometric projection matrix with element (i, j) equal to the probability that a photon pair produced in voxel j reaches the front faces of the detector pair i in the absence of attenuation and assuming perfect photon-pair colinearity. It incorporates a depth dependent geometric sensitivity that is calculated using the solid angle spanned by the voxel j at the faces of the detector pair i [1]. $\boldsymbol{P}_{det.blur}$ is the sinogram blurring matrix used to model photon pair non-colinearity, inter-crystal scatter and penetration [12], $\boldsymbol{P}_{attn}$ is a diagonal matrix containing the attenuation factors, and $\boldsymbol{P}_{det.sens}$ is again a diagonal matrix which contains the normalization factors that compensate for variations in detector pair sensitivity.

Most image priors used in PET image reconstruction have a Gibbs distribution of the form

$$p(\boldsymbol{x}) = \frac{1}{Z} \exp(-\beta U(\boldsymbol{x})) \tag{5}$$

where $U(\boldsymbol{x})$ is the energy function, β is the smoothing parameter that controls the resolution of the reconstructed image, and Z is the normalization constant or partition function. Combining the likelihood function and the image prior, the MAP reconstruction is found as:

$$\hat{\boldsymbol{x}}(\boldsymbol{y}) = \arg\max_{\boldsymbol{x} \geq 0} L(\boldsymbol{y}|\boldsymbol{x}) - \beta U(\boldsymbol{x}) \tag{6}$$

B. Approximations of Local Impulse Response and Covariance

The MAP estimator (6) is nonlinear in the data and its properties are object dependent. Therefore, we study the resolution and noise properties locally for each data set using the local impulse response and the covariance matrix.

The *local impulse response* for the jth voxel is defined as [3]

$$l^j(\hat{\boldsymbol{x}}) = \lim_{\delta \to 0} \frac{\mathcal{E}\hat{\boldsymbol{x}}(\boldsymbol{y}(\boldsymbol{x} + \delta \boldsymbol{e}_j)) - \mathcal{E}\hat{\boldsymbol{x}}(\boldsymbol{y}(\boldsymbol{x}))}{\delta} \tag{7}$$

---

[1] If $y_i < -2r_i < 0$, we set $y_i = -2r_i$.

where $\mathcal{E}$ denotes the expectation operator, $\hat{\boldsymbol{x}}(\boldsymbol{y})$ is the reconstruction from data $\boldsymbol{y}$, $\boldsymbol{y}(\boldsymbol{x})$ is the projection data from tracer distribution $\boldsymbol{x}$, and $\boldsymbol{e}_j$ is the jth unit vector.

Using a first order Taylor series approximation of (6) at the point $\boldsymbol{y} = \bar{\boldsymbol{y}}$ and the chain rule, one can derive the local impulse response for the MAP reconstruction at voxel j to be [3]

$$l^j(\hat{\boldsymbol{x}}) \approx [\boldsymbol{F} + \beta \boldsymbol{R}]^{-1} \boldsymbol{F} \boldsymbol{e}_j \tag{8}$$

and the covariance matrix [7]

$$Cov(\hat{\boldsymbol{x}}) \approx [\boldsymbol{F} + \beta \boldsymbol{R}]^{-1} \boldsymbol{F} [\boldsymbol{F} + \beta \boldsymbol{R}]^{-1} \tag{9}$$

where $\boldsymbol{F} \overset{\text{def}}{=} \boldsymbol{P}' D[1/\bar{y}_i] \boldsymbol{P}$ is the Fisher information matrix when using the Poisson likelihood model (1) or $\boldsymbol{F} \overset{\text{def}}{=} \boldsymbol{P}' D\left[1/((\boldsymbol{P}\boldsymbol{x})_i + s_i + 2r_i))\right] \boldsymbol{P}$ for the shifted-Poisson model (3). $D[x_i]$ represents a diagonal matrix with diagonal elements $x_i, i = 1, \ldots N$. $\boldsymbol{R}$ is the second derivative of the prior energy function $U(\hat{\boldsymbol{x}}(\bar{\boldsymbol{y}}))$. In the following, results are developed for the Poisson model only, extensions to the shifted-Poisson case are direct. Since (8) and (9) use derivatives of the log-likelihood and prior energy function up to order two only, they will be most accurate in cases where the objective function is locally quadratic.

Equations (8) and (9) both involve computation of the inverse of an $N \times N$ matrix where N is the number of image voxels. Even though one can avoid the computation of the matrix inverse by solving a set of linear equations for a voxel of interest [8], the computational cost can still be prohibitive for large numbers of voxels. Another problem is that the non-negativity constraint in (6) introduces nonlinearities that are not accounted for in the truncated Taylor series used to derive the approximations. This results in large errors in the variance estimate in low activity regions where the constraint is active. In the following section we develop approximations to (8) and (9) that are more readily computed. We also describe a method for modifying the covariances computed using (9) to account for the effect of the non-negativity constraint.

III. Resolution and Covariance for 3D PET

A. Simplified Expressions for Local Impulse Response and Covariance

In [9] we analyzed the resolution and covariance of MAP reconstructions for a simplified 2D PET system model using approximations similar to those in [7], [3], [13] including the assumption that the geometric response, $\boldsymbol{P}'_{geom} \boldsymbol{P}_{geom}$, is shift invariant. While this is a reasonable approximation in 2D, it is not applicable in 3D because of the "missing data" problem resulting from the finite number of detector rings. Here we extend the results in [9] to 3D by replacing the global invariance assumption with a local one. The idea of using a local invariance assumption in the context of shift-variant PET modeling was first proposed by Fessler and Booth [14] who applied this idea to developing fast preconditioners for conjugate gradient algorithms for optimization of cost functions similar to (6).

We can view the elements of the jth column of the Fisher information matrix $\boldsymbol{F}$ as representing an "image" associated with the jth voxel. We will assume that these Fisher information "images" vary smoothly as we move between the columns of $\boldsymbol{F}$ associated with neighboring voxels. We also assume that these images have local support, i.e. for the jth column, the significantly non-zero values are concentrated in the vicinity of the jth voxel. The rationale for these assumptions lies in the form: $\boldsymbol{F} \overset{\mathrm{def}}{=} \boldsymbol{P}' D[1/\bar{y}_i]\boldsymbol{P}$ (see for example the Fisher information matrix for a small scale problem shown in Figure 2 in [3]). We can then infer that the resolution and variance at voxel j is largely determined by the jth column of $\boldsymbol{F}$. Therefore, when estimating the resolution and variance at that voxel, we assume stationarity through out the scanner with the Fisher information matrix approximated by appropriate shifts of the elements of the jth column so that the resulting matrix $\boldsymbol{F}(j)$ has a block Toeplitz structure. This makes the computations in (8) and (9) tractable since a block Toeplitz matrix can be approximately diagonalized using a 3D FFT.

The Fisher information matrix must be positive semi-definite or equivalently its eigenvalues must be real and nonnegative. While the true $\boldsymbol{F}$ is guaranteed to have this property, the Toeplitz approximation may not. Consequently, we further modify the matrix by introducing the symmetry condition as follows. We first compute the jth column of $\boldsymbol{F}$ and arrange these values as a 3D image. For a $L \times L \times M$ voxel volume, we then shift this image so that the jth voxel is moved to the center voxel$(L/2+1, L/2+1, M/2+1)$. To ensure that the 3D FFT coefficients are real, we introduce the symmetry: $f(i,j,k) = \max\{f(i,j,k), f(L-i+1, L-j+1, M-k+1)\}$. Finally, we take the 3D FFT of the resulting image and truncate any negative coefficients to zero.

For a homogeneous prior with quadratic energy, $\boldsymbol{R}$ already has the block Toeplitz structure. However, if a spatially variant smoothing prior is used (see Section III-D), we can use a locally invariant approximation $\boldsymbol{R}(j)$ in a similar manner to that described above for $\boldsymbol{F}(j)$.

The local impulse response and covariance of voxel j can then be approximated by[2]

---

[2] When constructing a full covariance matrix using $Cov_j(\hat{\boldsymbol{x}})$ as the jth column, the resulting matrix may not be

$$l^j(\hat{\boldsymbol{x}}) \approx [\boldsymbol{F}(j)+\beta\boldsymbol{R}(j)]^{-1}\boldsymbol{F}(j)\boldsymbol{e}_j \tag{10}$$

$$Cov_j(\hat{\boldsymbol{x}}) \approx [\boldsymbol{F}(j)+\beta\boldsymbol{R}(j)]^{-1}\boldsymbol{F}(j)[\boldsymbol{F}(j)+\beta\boldsymbol{R}(j)]^{-1}\boldsymbol{e}_j \tag{11}$$

Since a block Toeplitz symmetric matrix is approximately block circulant, approximate inverses of $\boldsymbol{F}(j)$ and $\boldsymbol{R}(j)$ can be computed using a 3D Fourier transform.

Equations (10) and (11) can be used to evaluate the local impulse response and covariance at each voxel. The dominant computation cost is computing $\boldsymbol{F}(j)$ which involves one forward and one backprojection operation. If only a small number of voxels are of interest this approach is practical since the computational cost is similar to one reconstruction. However, evaluating these expressions for the whole image using (10) and (11) is prohibitive since the entire computation needs to be repeated for each voxel.

To study the local impulse response and variance throughout the field of view, we need to reduce the cost of computing $\boldsymbol{F}$. Using the factored system matrix (4), the Fisher information matrix $\boldsymbol{F}$ can be written as

$$\boldsymbol{F} = \boldsymbol{P}'_{geom}\boldsymbol{P}_{attn}\boldsymbol{P}'_{det.blur}\boldsymbol{P}_{det.sens}D\left[\frac{1}{\bar{y}_i}\right]\boldsymbol{P}_{det.sens}\boldsymbol{P}_{det.blur}\boldsymbol{P}_{attn}\boldsymbol{P}_{geom} \tag{12}$$

The approximations in [3] (Eq. (31)), [9] (Eq. (8)) and [14] (Eq. (13)) can not be used here because the computation is complicated by the spatially variant geometric and sinogram responses such that exact computation of the diagonal elements of $\boldsymbol{F}$ is impractical. To reduce the computation cost, we retain the shift variant components of the model but approximate $\boldsymbol{F}$ so that the time-consuming components of the computation are data independent and can be pre-computed and stored. In [1] we model the sinogram blurring, $\boldsymbol{P}_{det.blur}$, using a shift-variant local blurring kernel applied to the sinogram. This accounts for photon pair non-colinearity, inter-crystal scatter and crystal penetration. These effects can be decomposed into the following major components: (i) a projection shift due to crystal penetration; (ii) amplitude decrease of the local response due to detector blurring and (iii) a change in the shape of the local impulse response due to detector blurring. We therefore replace the approximations used in [3], [9] with the following approximation for (12) which explicitly incorporates the sinogram blurring factors:

$$\boldsymbol{F} \approx D[\kappa_j]D[\nu_j]^{-1}\boldsymbol{P}'_{geom}\boldsymbol{P}'_{det.blur}\boldsymbol{P}_{det.blur}\boldsymbol{P}_{geom}D[\nu_j]^{-1}D[\kappa_j] \tag{13}$$

symmetric because of the spatially variant system response. One can always obtain a symmetric covariance matrix by taking the average of the resulting matrix and its transpose.

where

$$\kappa_j \stackrel{\text{def}}{=} \sqrt{\frac{\nu_j^2 \sum_i g_{ij}^2 n_i^2 / [\boldsymbol{P}'_{det.blur}\bar{\boldsymbol{y}}]_i}{\sum_i g_{ij}^2}} \tag{14}$$

with g_{ij} the (i,j)th element of matrix $\boldsymbol{P}_{geom}$, n_i the (i,i)th element of the matrix product $\boldsymbol{P}_{det.sens}\boldsymbol{P}_{attn}$, and ν_j^2 the (j,j)th element of

$$\boldsymbol{P}'_{geom}\boldsymbol{P}'_{det.blur}\boldsymbol{P}_{det.blur}\boldsymbol{P}_{geom}$$

The κ_j^2 is an approximation of the (j,j)th diagonal element of $\boldsymbol{F}$ where the crystal penetration peak shift is accounted for by $\boldsymbol{P}'_{det.blur}\bar{\boldsymbol{y}}$. The decrease in the amplitude of the impulse response due to the detector blurring effect is approximated by the ratio $\nu_j^2/\sum_i g_{ij}^2$. The normalized spatially shape-variant impulse response in $\boldsymbol{F}$ is approximated using

$$D[\nu_j]^{-1}\boldsymbol{P}'_{geom}\boldsymbol{P}'_{det.blur}\boldsymbol{P}_{det.blur}\boldsymbol{P}_{geom}D[\nu_j]^{-1}$$

There is no optimality to the approximations (13) but we note that (13) is exact when $\boldsymbol{P}_{attn}$, $\boldsymbol{P}_{det.sens}$ and $D\left[\frac{1}{\bar{y}_i}\right]$ are all equal to the identity matrix.

Using this approximation, $\boldsymbol{P}'_{geom}\boldsymbol{P}'_{det.blur}\boldsymbol{P}_{det.blur}\boldsymbol{P}_{geom}$ becomes the dominant computation load in computing $\boldsymbol{F}$. Since it is independent of the data, it can be precomputed. Furthermore, by taking advantage of the rotational symmetry of the PET system, we need only compute the columns that correspond to the voxels in a single plane containing the symmetry axis of the scanner. We refer to voxels in this plane as "base voxels". All the other columns can be approximated using linear or nearest neighbor interpolation. This reduces the computation time and storage space required for $\boldsymbol{P}'_{geom}\boldsymbol{P}'_{det.blur}\boldsymbol{P}_{det.blur}\boldsymbol{P}_{geom}$ to a practical level.

We can now write the local impulse response (10) and covariance (11) in Fourier transform form as:

$$\begin{aligned}
l^j(\hat{\boldsymbol{x}}) &\approx \boldsymbol{Q}'\boldsymbol{Q}[\boldsymbol{F}(j)+\beta\boldsymbol{R}(j)]^{-1}\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{F}(j)\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{e}_j \\
&\approx \boldsymbol{Q}'[\boldsymbol{Q}\boldsymbol{F}(j)\boldsymbol{Q}'+\beta\boldsymbol{Q}\boldsymbol{R}(j)\boldsymbol{Q}']^{-1}[\boldsymbol{Q}\boldsymbol{F}(j)\boldsymbol{Q}']\boldsymbol{Q}\boldsymbol{e}_j \\
&\approx \boldsymbol{Q}'D[\frac{\lambda_i(j)}{\lambda_i(j)+\beta\kappa_j^{-2}\mu_i(j)}]\boldsymbol{Q}\boldsymbol{e}_j \qquad (15)\\
Cov_j(\hat{\boldsymbol{x}}) &\approx \kappa_j^{-2}\boldsymbol{Q}'D[\frac{\lambda_i(j)}{(\lambda_i(j)+\beta\kappa_j^{-2}\mu_i(j))^2}]\boldsymbol{Q}\boldsymbol{e}_j \qquad (16)
\end{aligned}$$

where $\{\lambda_i(j),\ i=1,\ldots,N\}$ is the 3D Fourier transform of the positive semi-definite approximation of the central column of the block-Toeplitz matrix formed from the jth column of

$$D[\nu_j]^{-1}\boldsymbol{P}'_{geom}\boldsymbol{P}'_{det.blur}\boldsymbol{P}_{det.blur}\boldsymbol{P}_{geom}D[\nu_j]^{-1},$$

and $\{\mu_i(j),\ i = 1, \ldots, N\}$ is the 3D Fourier transform of the central column of $\boldsymbol{R}$. $\boldsymbol{Q}$ and $\boldsymbol{Q}'$ represent the Kronecker form of the 3D Fourier transform and its inverse, respectively.

For space invariant priors with quadratic energy functions, (15) and (16) can be simplified to

$$l^j(\hat{\boldsymbol{x}}) \approx \boldsymbol{Q}' D[\frac{\lambda_i(j)}{\lambda_i(j) + \beta\kappa_j^{-2}\mu_i}]\boldsymbol{Q}\boldsymbol{e}_j, \tag{17}$$

$$Cov_j(\hat{\boldsymbol{x}}) \approx \kappa_j^{-2}\boldsymbol{Q}' D[\frac{\lambda_i(j)}{(\lambda_i(j) + \beta\kappa_j^{-2}\mu_i)^2}]\boldsymbol{Q}\boldsymbol{e}_j \tag{18}$$

where the μ_i's are the 3D Fourier transform of the central column of $\boldsymbol{R}$.

B. CRC and Variance

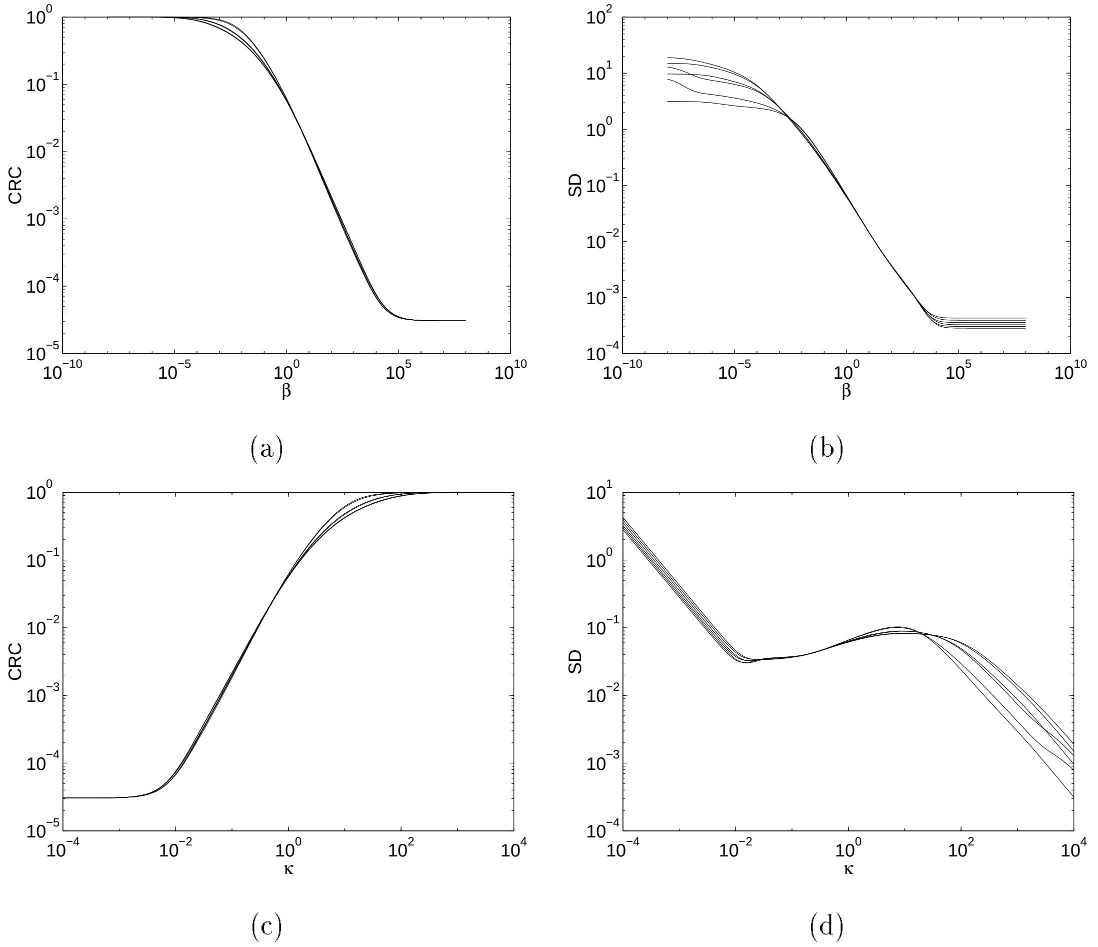
We can reduce (17) and (18) to scalar measures by considering only the variance and the local contrast recovery coefficient (CRC) which we define as $crc_j = l_j^j(\hat{\boldsymbol{x}})$. The CRC can be used as an alternative to the full width at half maximum (FWHM) as a measure of resolution which has the advantage that it can be directly computed from (17) (we will examine the relationship between CRC and FWHM in Section IV-D). The CRC and variance for the jth voxel are given by

$$crc_j \approx \frac{1}{N}\sum_{i=0}^{N-1}\frac{\lambda_i(j)}{\lambda_i(j) + \beta\kappa_j^{-2}\mu_i} \tag{19}$$

$$var_j \approx \kappa_j^{-2}\frac{1}{N}\sum_{i=0}^{N-1}\frac{\lambda_i(j)}{(\lambda_i(j) + \beta\kappa_j^{-2}\mu_i)^2} \tag{20}$$

Expressions (19) and (20) provide direct insight into the spatially variant properties of MAP reconstructions: since the only function of the data is the quantity κ_j, we can, in the absence of any data, determine resolution and noise properties at each voxel as a function of κ_j. The spatial variations in the κ_j's associated with a given source distribution imply spatial variation in resolution and variance. The β value necessary to achieve a desired CRC or variance can then be chosen once κ_j has been computed as we describe in Section III-D. The results (19) and (20) require that the mean of the data is available to compute κ_j. However they can also be used in a "plug-in" mode in which experimental data is used to estimate κ_j. This issue is addressed in Section V.

We can evaluate (19) and (20) to show the dependence of the CRC and variance on the hyperparameter β and κ_j. The PET system simulated here was the microPET system [15] with 8 image planes of 64×64 voxels. We used a second order (26 neighbors) 3D prior with a quadratic energy function. Because of the circular symmetry of the PET system, we need only



(a) (b)

(c) (d)

Fig. 1. CRC and SD (standard deviation) curves for 6 different locations in the outermost and central axial planes: (a) CRC vs. β ($\kappa = 1$) (b) SD vs. β ($\kappa = 1$) (c) CRC vs. κ ($\beta = 1$) and (d) SD vs. κ ($\beta = 1$).

consider radial and axial variations in CRC and variance. We selected 3 points with different radial positions in both the outermost and the central transaxial planes. The results are shown in Fig. 1. Each plot is similar to that shown for a 2D PET system in [9] with two inflection points. Note that in Fig. 1(d) there is a range of κ values in which the standard deviation varies slowly (although these curves are less flat than their 2D equivalents in [9]) indicating a range of values over which the image standard deviation will vary very little. This observation is confirmed in our simulation studies below.

C. Compensation for Non-negativity Constraints

The development of (16) in Section II-B is based on a first order Taylor series approximation and cannot account for the non-negativity constraint typically used in MAP reconstruction. This results in large errors in covariance estimates for low activity regions [7], [9]. In this section we develop a method to modify the preceding results to account for the effect of this constraint.

We first consider the effect of the constraint on the voxel-wise variance. We assume that if the non-negativity constraint were not imposed, the voxel intensities of the MAP reconstructions, conditioned on the true image, would be Gaussian random variables. Empirical evidence supporting this assumption is provided later. We further assume that the effect of the non-negativity constraint is to modify this Gaussian distribution by replacing all negative voxel values with zero, i.e. the constraint truncates the original Gaussian distribution in the negative range, but does not change the distribution of the voxel values in the positive range. Under this assumption, the actual distribution of the voxel values will be a "truncated Gaussian" with probability density function

$$p(x) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} & \text{if } x > 0 \\ \delta(x)\left[\frac{1}{2} - \frac{\sqrt{\pi}}{4}\mathrm{erf}(\frac{\mu}{\sqrt{2\sigma^2}})\right] & \text{if } x = 0 \\ 0 & \text{if } x < 0 \end{cases} \qquad (21)$$

where μ and σ^2 are the mean and variance of the original Gaussian distribution, respectively, $\mathrm{erf}(x)$ is the error function and $\delta(x)$ is the Dirac delta function.

Because of the truncation at $x = 0$, the actual mean, μ_x, and variance, σ_x^2, of the truncated Gaussian distribution differ from the original mean, μ, and variance, σ^2, and are given by:

$$\mu_x(\mu,\sigma) = \sqrt{\frac{\sigma^2}{2\pi}} e^{-\frac{m^2}{2\sigma^2}} + \frac{\mu}{2}\left[1 + \mathrm{erf}(\frac{\mu}{\sqrt{2\sigma^2}})\right] \qquad (22)$$

$$\sigma_x^2(\mu,\sigma) = \mu\sqrt{\frac{\sigma^2}{2\pi}} e^{-\frac{m^2}{2\sigma^2}} + \frac{1}{2}(\mu^2+\sigma^2)\left[1 + \mathrm{erf}(\frac{\mu}{\sqrt{2\sigma^2}})\right] - [\mu_x(\mu,\sigma)]^2 \qquad (23)$$

It is straightforward to show that

$$\frac{\mu_x(\mu,\sigma)}{\sigma} = f(\frac{\mu}{\sigma}) \qquad (24)$$

where

$$f(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} + \frac{\xi}{2}\left[1 + \mathrm{erf}(\frac{\xi}{\sqrt{2}})\right] \qquad (25)$$

and

$$\frac{\sigma_x^2(\mu,\sigma)}{\sigma^2} = g(\frac{\mu}{\sigma}) \qquad (26)$$

where

$$g(\xi)=\frac{\xi}{\sqrt{2\pi}}e^{-\frac{\xi^2}{2}}+\frac{1}{2}(\xi^2+1)\left[1+\mathrm{erf}(\frac{\xi}{\sqrt{2}})\right]-f(\xi)^2 \tag{27}$$

i.e. $\frac{\mu_x(\mu,\sigma)}{\sigma}$ and $\frac{\sigma_x^2(\mu,\sigma)}{\sigma^2}$ are both functions of $\frac{\mu}{\sigma}$. Therefore, if we can find $\frac{\mu_x(\mu,\sigma)}{\sigma}$, we will be able to calculate $\frac{\sigma_x^2(\mu,\sigma)}{\sigma^2}$.

Since (20) implicitly assumes an unconstrained reconstruction, the variance from (20) is an estimate of the original Gaussian variance, σ^2. To account for the effect of the constraint, we need to replace this variance with σ_x^2, or equivalently compute the ratio σ_x^2/σ^2. The mean of the truncated Gaussian distribution, μ_x, is actually the mean of the corresponding voxel in the MAP reconstructions. This can be estimated as the ensemble mean of a set of Monte Carlo reconstructions, or approximated by reconstruction of the noiseless projection data [7], [6]. With this mean, μ_x, and the unconstrained variance σ, we can invert (24) to find μ/σ. We can then compute the fraction σ_x^2/σ^2 using (26). These computations can be performed rapidly using look-up tables for (25) and (27).

In practical situations, where neither the noiseless projection data nor sufficient number of independent data sets are available, a single noisy MAP reconstruction may be the only source that can be used to estimate μ_x. If so, the noise in the reconstruction will affect the accuracy of the estimate. As a result, an over-smoothed MAP reconstruction may be more suitable for the purpose of computing μ_x than the original reconstruction.

After we obtain the voxel-wise variance, σ_x^2, under the non-negativity constraint, we approximate the image covariance matrix by

$$Cov \approx D[\sigma_x(j)]Corr D[\sigma_x(j)] \tag{28}$$

where $Corr$ is the correlation matrix estimated from (18). This approximation, in which the correlation and variance terms are decoupled, is similar to that used in [19] for computing the variance of regions of interest. It is also similar in spirit to the approximation of the Fisher information matrix used in Section III-A.

An implicit assumption in this variance-compensation technique is that the non-negativity constraint affects each voxel independently. In practice, the impact of the smoothing prior is to couple the voxels so that activation of the constraint at one voxel will affect the variance of its neighbors. This will affect the accuracy of the approximation. However as we show in Section IV, the approximations appear reasonably accurate and are significant improvements over previously

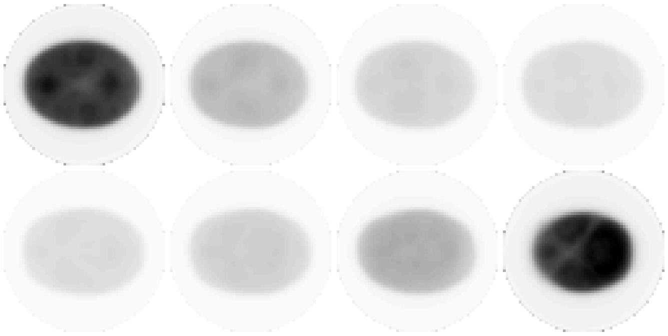


Fig. 2. The values of κ_j^{-2} displayed as an image for a simulated scaled 3D Hoffman brain phantom for the microPET system configuration. The image is 8 planes of 64×64 voxels. An inverse gray scale is used for better visualization of spatial variations.

reported results in which the non-negativity constraint was ignored [7], [9].

D. Uniform Resolution Reconstruction

From Fig. 1(c) we see that the CRC of the MAP reconstruction with a constant β is highly dependent on κ_j. Although κ_j generally changes smoothly inside the support of the object, there is still substantial variation from the center to the axial boundary, as shown in Fig. 2. This causes the CRCs and hence resolution in 3D PET to be highly non-uniform. In some situations it may be desirable to reconstruct images with uniform CRCs. For instance, when multi-bed acquisitions are overlapped in the axial direction, the variance at the axial boundary of each bed position can be reduced by adding together reconstructions from overlapped planes that correspond to the same position. If resolution is mismatched this may produce artifacts. Uniform axial resolution in the form of matched CRCs may avoid this problem.

In order to achieve uniform contrast recovery, the hyperparameter β must be spatially variant. For any desired CRC (between 0 and 1), we can find the corresponding $\beta_j\kappa_j^{-2} = \eta_j^*$ for each voxel j using (19). Because (19) as a function of $\eta = \beta_j\kappa_j^{-2}$ can be precomputed for all the base voxels, the η_j^* for the desired resolution can be found, independently of the data, using a lookup table. Given estimates of the κ_j, we then set $\beta_j = \eta_j^*\kappa_j^2$. This method is straightforward but fails to account for the fact that the β's are being varied throughout the volume. The spatial variation in β introduces a local interaction effect so that the look-up table approach does *not* produce uniform CRCs. The effect is particularly pronounced towards the edge of the axial field of view where the κ_j (and hence β_j) values can change significantly from one plane to the next.

To obtain uniform CRCs it is necessary to solve a coupled system of equations. When we vary the smoothing parameters throughout the image, we assign a separate β_j to each voxel and redefine the energy function as:

$$U(\boldsymbol{x}) = \frac{1}{2}\sum_{j=1}^{N}\sum_{k\in N_j, k>j} \rho_{jk}\sqrt{\beta_j\beta_k}(x_j - x_k)^2 \tag{29}$$

where ρ_{jk} is the reciprocal of the Euclidean distance between voxel j and k. The second derivative of (29) is

$$\boldsymbol{R}(j,k) = \begin{cases} -\rho_{jk}\sqrt{\beta_j\beta_k} & \text{if } j \neq k,\ k \in N_j \\ \sum_{l\in N_j}\rho_{jk}\sqrt{\beta_j\beta_k} & \text{if } j = k \end{cases} \tag{30}$$

For exact uniform resolution, we would need to iterate between computing the Fourier transform coefficients $\mu_i(j)$ of the symmetric Toeplitz approximation $\boldsymbol{R}(j)$ and updating the β_j's using (19) with the new $\mu_i(j)$'s. This is a very computationally intensive procedure and probably not warranted since (19) is only an approximation. A more practical solution is to consider only the diagonal elements of $\boldsymbol{R}$ and solve the following set of equations[3]

$$\sum_{l\in N_j}\rho_{jl}\sqrt{\beta_j\beta_l} = \eta_j^*\kappa_j^2\sum_{l\in N_j}\rho_{jl} \quad \forall j \tag{31}$$

Eq. (31) may not have an exact solution but can be solved iteratively in a least squares sense using an iterative coordinate descent method to minimize the error function

$$E = \sum_j\left(\sum_{l\in N_j}\rho_{jl}\sqrt{\beta_j\beta_l} - \eta_j^*\kappa_j^2\sum_{l\in N_j}\rho_{jl}\right)^2 \tag{32}$$

We have found that a coordinate-wise descent algorithm converges rapidly, taking a small fraction of the image reconstruction time for the microPET system simulated here.

The following scheme can be used for reconstructing uniform CRC images with quadratic priors:

1. select a desired CRC.
2. for each voxel j, use a lookup table to find the corresponded η_j^* for the given CRC.
3. compute the κ_j's using (14) and the mean of the PET data (or actual PET data when used in "plug-in" mode).
4. use a coordinate descent algorithm to find the β_j's that minimize (32).
5. reconstruct the image with the spatially variant smoothing parameters β_j.

---

[3] This is equivalent to approximating matrix $\boldsymbol{R}$ by $D[r_j]\boldsymbol{R}_0D[r_j]$, where $r_j = (\sum_{l\in N_j}\rho_{jl}\sqrt{\beta_j\beta_l}/\sum_{l\in N_j}\rho_{jl})^{-1/2}$ and $\boldsymbol{R}_0$ is the second derivative of the homogeneous quadratic energy function.

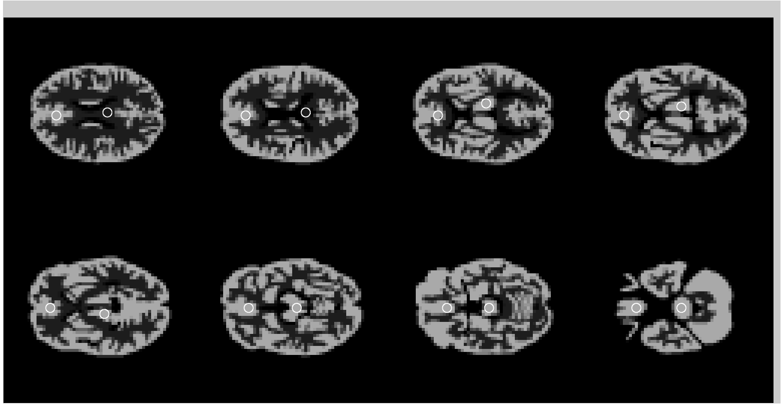


Fig. 3. The 3D Hoffman brain phantom used in our Monte Carlo studies. The white circles indicate the voxels selected for evaluation of the CRC approximation.

IV. Monte Carlo Validation

We used computer Monte Carlo simulations to evaluate the approximations described above. All simulations were based on the geometry of the microPET scanner [15] which consists of 8 rings with 240 2mm x 2mm x 10mm LSO detectors in each ring. The field of view is 112mm transaxially by 18mm axially and all images were reconstructed on 8 2.25mm thick planes with 64×64 1.5mm voxels. Data were generated using forward projection through the factored matrix model developed in [1] which includes a spatially varying geometric response $\boldsymbol{P}_{geom}$ and detector response blurring kernels $\boldsymbol{P}_{det.blur}$. The latter were computed using Monte Carlo modeling of photon-pair production and interaction within the detector ring. The phantom image was a scaled 3D digital Hoffman brain [20] as shown in Fig. 3. The normalization factors were based on measurements from a cylindrical normalization source collected in the microPET scanner. The attenuation correction factors were computed analytically assuming a constant attenuation coefficient 0.095cm^{-1} throughout the support of the phantom. The average number of counts in each data set was 6M and included a 10% uniform scatter background. All the images were reconstructed using 60 iterations of a non-negatively constrained preconditioned conjugate gradient (PCG) algorithm with a second order quadratic energy function as described in [2].

A. Statistical Distribution of Image Voxel Values

In developing the variance approximation that accounts for the effect of the non-negativity constraint, we assumed a truncated Gaussian as described in Section III-C. To investigate this conjecture, we calculated the sample distribution for individual voxels in Monte Carlo recon-

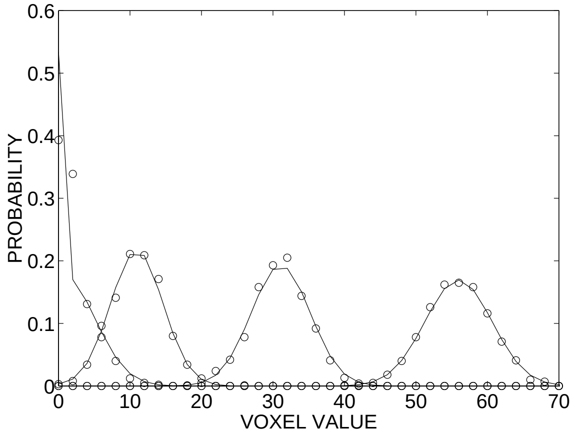
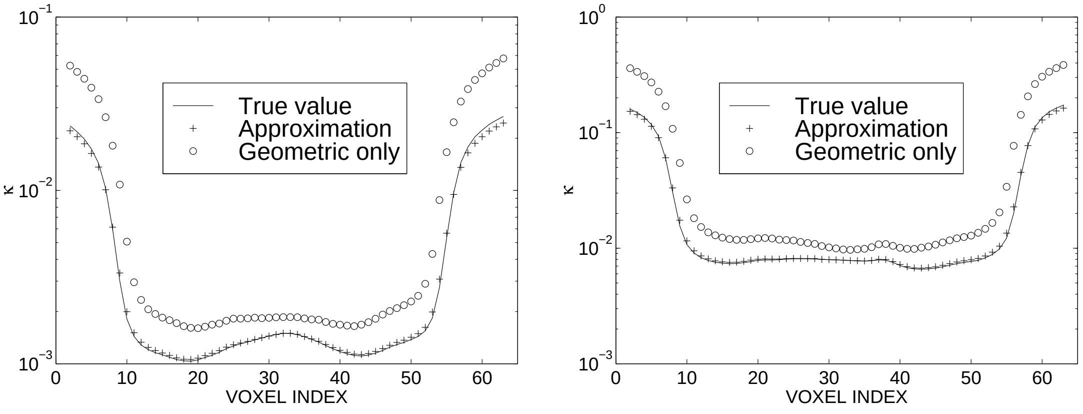


Fig. 4. Intensity distribution of image voxel values for four points: one CSF, one white matter and one gray matter, and one grey-white partial volume. The circles represent the histogram of voxel values of 1000 Monte Carlo reconstructions. The solid lines represent the estimated probabilities in each histogram bin using the truncated Gaussian distribution model.

structions of the brain phantom. Four points of interest were selected: one each in CSF, white and gray matter and one in a gray-white partial volume voxel. The sample distributions, overlaid with truncated Gaussian distributions based on the Monte Carlo sample mean and variance are shown in Fig. 4. There is generally a good match between the sample histogram and the truncated Gaussian distributions. However, the truncated Gaussian density tends to overestimate the probability of the voxel values being zero while underestimating the probability of occupying the neighboring histogram bin. This could result in underestimation of the variance in low intensity regions, which we investigate further in Section IV-E.

B. Approximation of κ_j

The κ_j values for the phantom were computed using (14). If the approximation of the Fisher information matrix (12) were exact, then the κ_j^2's would represent the diagonal elements of the Fisher information matrix. These can be computed exactly from $\boldsymbol{F} = \boldsymbol{P}' D[1/\bar{y}_i]\boldsymbol{P}$. Fig. 5 shows profiles through the image of κ_j values that pass through the symmetry axis of the scanner for the first and central transaxial planes. Also shown are the values that would be computed if the sinogram blurring factors are dropped from the approximation (denoted "geometric only" in the figure). This figure demonstrates very little loss in accuracy in κ_j as a result of the approximation and also that inclusion of the sinogram blurring factors is important for an accurate



(a) (b)

Fig. 5. Comparison of the κ values computed using (14) ("approximation") with the true values of the diagonals of the Fisher information matrix ("true value"): (a) the first transaxial plane; (b) the central transaxial plane. The "geometric only" values represent the estimates when the sinogram blurring factors are dropped from (14).

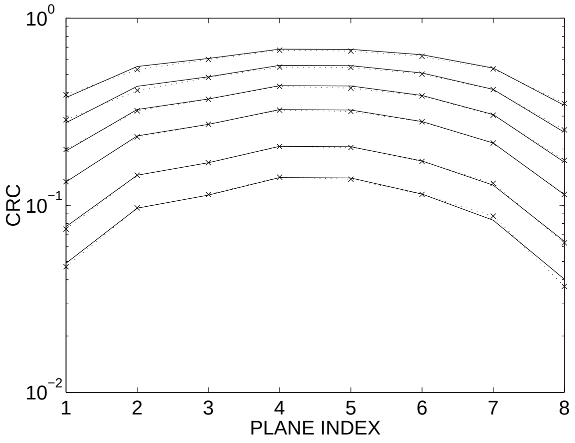
approximation.

C. Approximation of CRCs

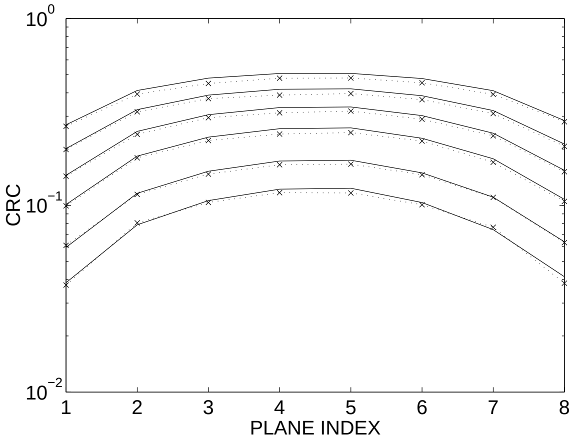
We selected two points of interest in each image plane at which to evaluate the CRC approximation (19); these are indicated in Fig. 3. The "ground truth" CRC was calculated from reconstructions from two noiseless data sets: (i) the original phantom sinogram, and (ii) the sinogram of the phantom after adding a perturbation at the point of interest. The approximations were computed using (19). In both cases, a quadratic energy function with a second order neighborhood was used. Fig. 6(a) shows the CRC values for voxels lying approximately along the symmetry axis of the scanner. Each curve corresponds to a different smoothing parameter β, ranging from 2.5×10^{-5} (top) to 0.001 (bottom). The approximation shows an almost exact match with the "ground truth" values. In Fig. 6(b) we show the CRC values for off-center voxels for the same range of β values. In this case there is a small increase in the error but they are at most a few percent.

D. CRC vs FWHM

As we discussed previously, we characterize resolution through the local CRC rather than the traditional FWHM resolution. To achieve some insight into the relationship between these, we



(a)



(b)

Fig. 6. The CRCs computed using the approximation (19) compared with ground truth values. (a) Comparison for the voxels close to the symmetry axis of the scanner as indicated in Fig. 3; (b) comparison for off-axis voxels also shown in Fig. 3. The solid lines denote the approximation results and ×'s denote the measured ground truth.

computed the FWHM of the local impulse response at each of the locations studied in Section IV-C. The local impulse response is not symmetric so we computed a mean FWHM in the transaxial plane using

$$mean\ FWHM = \sqrt{\frac{\text{area of the contour at half maximum}}{4\pi}}.$$

The FWHM versus CRC curves are plotted in Fig. 7. This figure indicates a monotonic relationship between FWHM and CRC for each voxel with very similar curves for voxels at a fixed radial distance from the scanner axis. However the height of these curves vary with radial distance

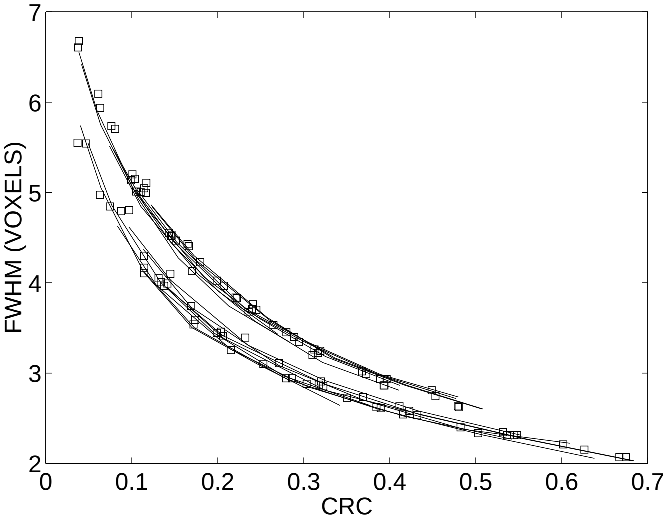


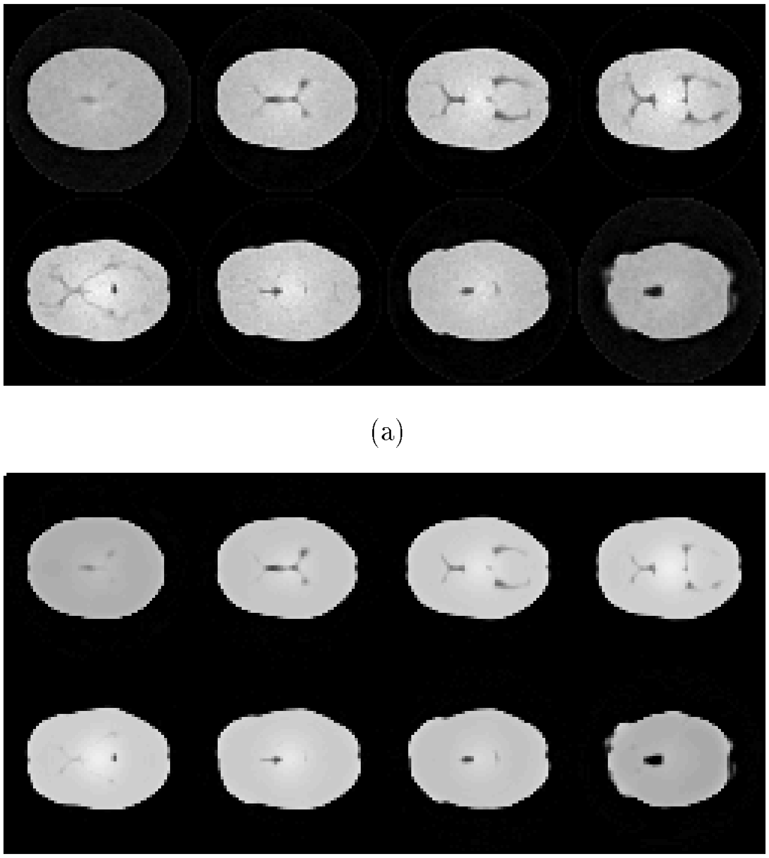
Fig. 7. Relation between FWHM and CRC. Squares denote ground truth and solid lines the theoretical approximation. The lower set of curves correspond to the points around the central axis of the scanner indicated in Fig. 3 while the upper set correspond to the points in the same figure that are off axis. Different points on the FWHM vs. CRC curve were generated using different values of β.

and consequently we cannot claim that a constant CRC throughout the volume translates to a constant FWHM, or vise versa. We note that the asymmetry of the local impulse response indicates that *any* scalar measure of resolution at a point will be deficient in characterizing the response, and for our purposes the CRC has distinct advantages over FWHM in terms of our ability to directly compute it.

E. Approximation of Variance

To investigate the accuracy of the approximate variance expression (20) we computed the voxel-wise variances from 1000 independent reconstructions of the phantom and compared these with the values computed using (20). Fig. 8 shows the standard deviation images for both the Monte Carlo results and the theoretical approximations. A selected profile passing through the CSF region in the second plane is shown in Fig. 9. The theoretical approximations are generally in good agreement with the Monte Carlo results.

We illustrate the impact of the method in Section III-C for compensating for the effect of the non-negativity constraint in Fig. 10. In the left part of the figure, we show a scatter plot of the uncorrected standard deviation (computed using (20)) versus the Monte Carlo standard deviations. On the right, we show the corrected standard deviations versus the Monte Carlo



(a)

(b)

Fig. 8. Standard deviation images computed using (a) the Monte Carlo method from 1000 reconstructions, and (b) the theoretical approximation (20). The order of the image planes are from left to right, top row: plane 1 (upper axial edge) to 4 (center); bottom row: plane 5 (center) to 8 (lower axial edge).

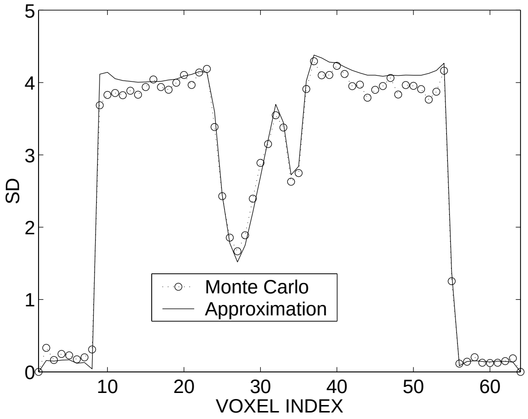
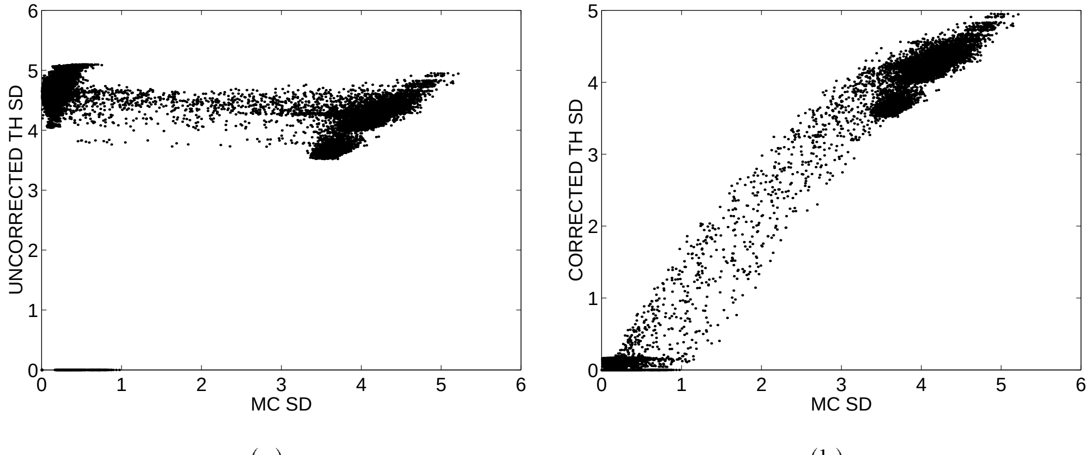


Fig. 9. Comparison of center transaxial profiles passing through CSF region in the second plane of standard deviation images in Fig. 8.



(a) (b)

Fig. 10. Scatter plots of variance estimates: (a) uncorrected theoretical standard deviations (the standard deviations estimate from (20)) vs. Monte Carlo standard deviations; (b) corrected theoretical standard deviations vs. Monte Carlo standard deviations

results. The scatter plot shows a tendency to underestimate the variance, which is consistent with the observation in Section IV-A. However, the results are a substantial improvement on those that do not compensate for the effect of the non-negativity constraint.

The remaining differences between the Monte Carlo and theoretical variances are due to a combination of factors: residual variance in the Monte Carlo sample statistics, deviations of the MAP images from the assumed truncated-Gaussian model, and errors caused by the local stationary approximation. Since our variance computation scheme is based on a sequence of approximations, it is not surprising that the computed variances are not exact. However, we anticipate that these variances are sufficiently accurate to be of practical value (we will quantify the accuracy of the approximation in the next section).

F. Estimating the variance of integrated ROI activity

One of the important applications of covariance estimation is to compute the uncertainty in ROI quantitation. Here we use the theoretical covariance expression (17) to estimate the variance of the integrated activity in several ROIs. The results are then compared to the variances estimated using the Monte Carlo method with 1000 independent reconstructions. Sixty-five ROI centers in the phantom were selected. For each ROI center, we drew 8 concentric circular regions with radius varying from 1 to 8 voxels, so there are totally 65×8=520 ROIs.

For each ROI, the total mean activity, R, was computed as

$$R = \frac{1}{\sum_{j=1}^{N} f_j} \sum_{j=1}^{N} f_j x_j \tag{33}$$

where $\{f_j,\ j = 1, \ldots, N\}$ is an indicator function for the jth ROI. The variance of R is then

$$Var(R) = \frac{1}{(\sum_{j=1}^{N} f_j)^2} \sum_{j=1}^{N} \sum_{k=1}^{N} f_j f_k \sigma_x(j) \sigma_x(k) Corr(x_j, x_k) \tag{34}$$

where $\sigma_x(j)$ denotes the estimated variance of voxel j with compensation for the effect of the non-negativity constraint and $Corr(x_j, x_k)$ is the correlation between voxels j and k. Substituting (18) in (34) and assuming that $Corr(x_j, x_k)$ is stationary within the ROI, we get

$$Var(R) = \frac{\kappa_j^{-2}}{(\sum_{j=1}^{N} f_j)^2} \sum_{i=1}^{N} \frac{|F_i|^2 \lambda_i}{(\lambda_i + \beta\kappa^{-2}\mu_i)^2} / \sum_{i=1}^{N} \frac{\lambda_i}{(\lambda_i + \beta\kappa^{-2}\mu_i)^2} \tag{35}$$

where $\{F_i,\ i = 1, \ldots, N\}$ is the 3D Fourier transform of $\{f_j \sigma_x(j),\ j = 1, \ldots, N\}$.

As a comparison, the ratio of the Monte Carlo standard deviation to the theoretical estimate is plotted as a function of the theoretical value in Fig. 11. For most ROIs, the ratio lies in the range 0.95 to 1.05, with the largest relative error in all ROIs of 8%.

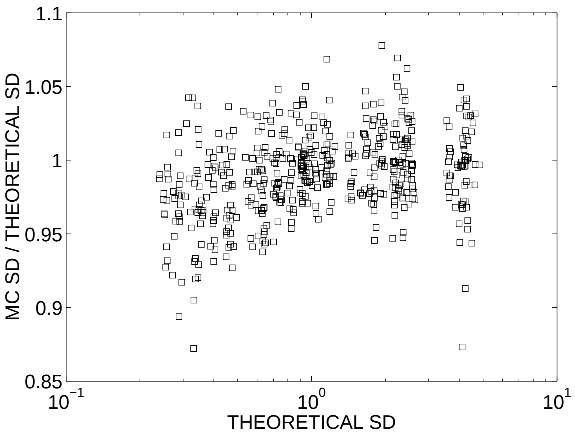


Fig. 11. Ratio of Monte Carlo standard deviation estimates to the theoretical results vs. the theoretical standard devation for the ROI quantitation study.

To quantify the accuracy of the approximation, we calculated the root mean squared error (RMSE) between the Monte Carlo results and the theoretical approximations

$$\text{RMSE} = \sqrt{\frac{1}{N_r} \sum_{i=1}^{N_r} \left(\frac{Var_i^{mc} - Var_i^{app}}{Var_i^{mc}} \right)^2}$$

where N_r is the number of ROIs. The resulting RMSE was 6.4%. Frieden [22] shows that for a Gaussian random variable, the relative error of the estimate of the variance with L i.i.d. samples is $\sqrt{2/(L-1)}$. For $L = 1000$ we would expect an error of approximately 4.5% in the Monte Carlo study. Thus the accuracy of the theoretical approximation is comparable to that of the Monte Carlo estimate with 1000 samples. Since the computational cost of this approximation is less than that for one reconstruction, the advantage in computation time is quite significant. Possibly more importantly, the theoretical approximation allows estimation of the variance of individual reconstructions using a single noisy measurement using a "plug-in" form of the approximate variance (see Section V).

G. Uniform CRC Reconstruction

Fig. 6 clearly shows spatially variant CRCs when spatially invariant smoothing priors are used. The resolution changes substantially in the axial direction for both center and off-center voxels. Here we demonstrate using the spatially variant smoothing prior developed in Section III-D to reconstruct near uniform resolution images. We selected a desired CRC of 0.3. A coordinate descent algorithm for minimizing (32) to select the appropriate β_j's took only 10 iterations to effectively converge. Fig. 12 shows the measured CRCs of the images reconstructed using the spatially variant β_j's computed to give a uniform CRC of 0.3. The CRCs were very close to the desired value for all planes and near uniform in both axial and transaxial directions. The theoretically predicted CRC (i.e. those computed using the approximate expression (19)) are also shown in Fig. 12. These are slightly more uniform than those based on the measured CRCs. This would appear to indicate that the remaining source of non-uniformity lies in the errors in the approximation of the system response rather than the manner in which the spatially variant smoothing parameters are computed. However, when we chose the β_j's without solving (31) but instead using a lookup table, as suggested in [3], the CRCs at the boundary planes dropped to approximately 0.25 (a 16% error) due to oversmoothing from adjacent planes. Thus, for uniform CRCs throughout the field of view, it is necessary to consider the coupling effect between spatially variant β's as is done in (31).

We also investigated the variance distribution for the uniform resolution reconstructions. The voxel-wise variances were computed using both the Monte Carlo method from 1000 independent reconstructions and the theoretical approximation. Comparisons of the standard deviation images and selected profiles are shown in Fig. 13 and Fig. 14, respectively. As would be expected,

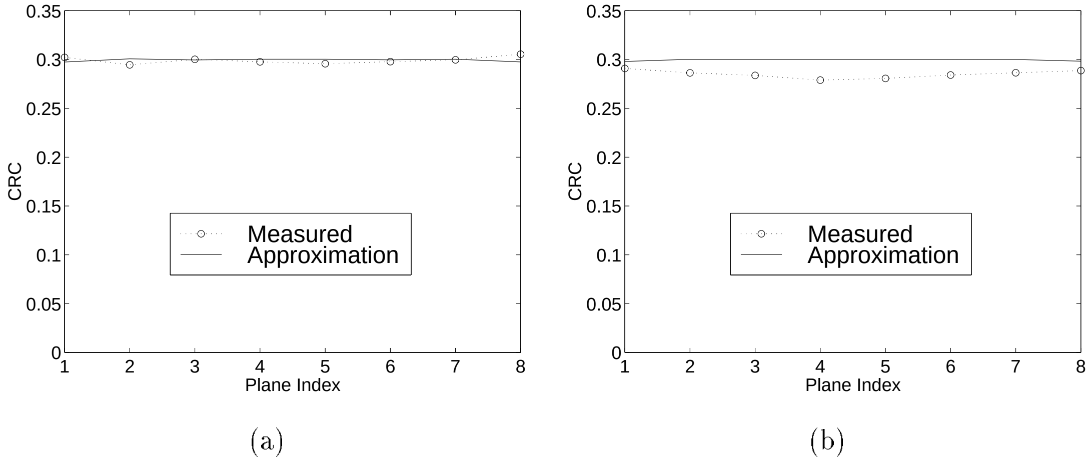


Fig. 12. Uniform CRC reconstruction with spatially varying β: the CRCs of (a) voxels along the central axis of the scanner and (b) off-center voxels. Measured values were those obtained by computing the CRC by perturbing the phantom at the point of interest and measuring the resulting change in intensity in the reconstruction at that point. Approximate values were computed using (19).

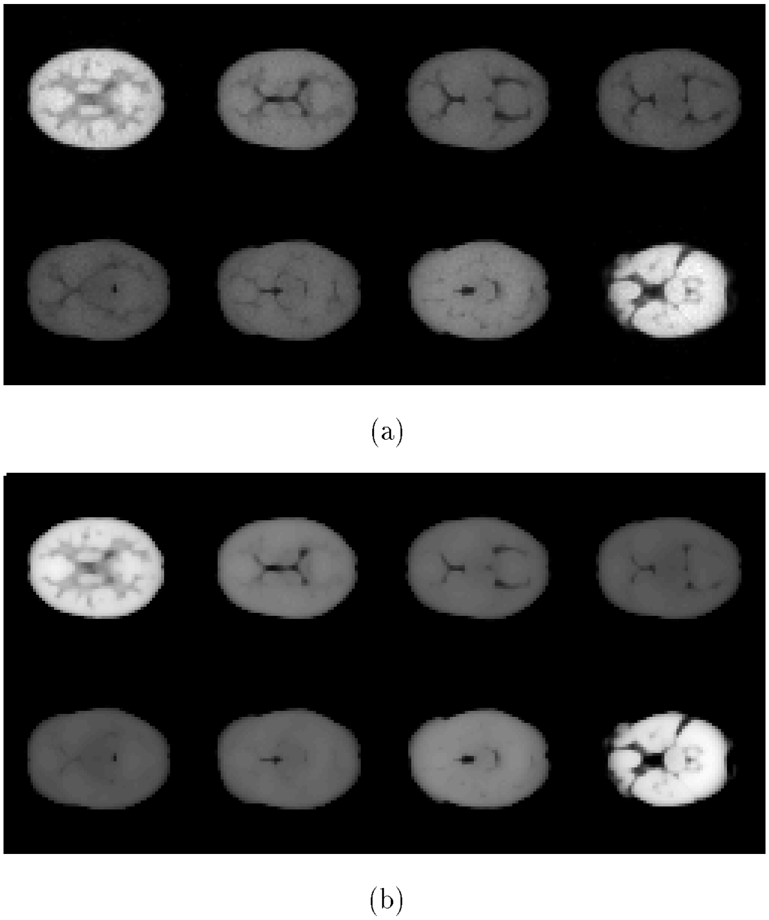
the variances at the axial boundaries are increased because of increased contrast recovery. The theoretical results are in good agreement with the Monte Carlo results, again demonstrating the effectiveness of the theoretical approximations.

V. Validation With Monkey Brain Phantom Scans

To investigate the effectiveness of the covariance approximation in plug-in mode, we used experimental data collected from a baby monkey brain phantom scanned using the microPET scanner [15]. Forty-one equal count data sets were recorded with each data set having about 6 million events. The 41 data sets were reconstructed using PCG MAP with $\beta = 0.0002$. Examples of the reconstructions for a single data set are shown in Fig. 15.

A. Covariance Computation using a Modified Plug-in Method

Since we do not have noise-free data, we must use the noisy data to compute the κ_j values in (14) which we use in turn to compute (18). One can use the direct "plug in" method in which the measured data is directly used in place of the mean [7]. However this will result in a biased estimate since we are taking the reciprocal of the data value as an estimate of the reciprocal of its mean. It is well known and illustrated in Fig. 16 that $E[1/y] \neq 1/E[y]$ for a Poisson random variable y. This direct plug in method produces a small positive bias for large $E[y]$. The negative



(a)

(b)

Fig. 13. Standard deviation images for the uniform resolution reconstructions. Computed using (a) the Monte Carlo method, and (b) the theoretical approximation. The order of the image planes are from left to right, top row: plane 1 (upper axial edge) to 4 (center); bottom row: plane 5 (center) to 8 (lower axial edge).

bias when $E[y]$ is small is due to the computation of the reciprocal of $[1/y]$ as 1 when $y = 0$.

Bias could be reduced by forward projecting a reconstructed image. However, we would often like to compute the variance before reconstruction. We therefore use the following correction method. We first note that for a Poisson random variable y:

$$\begin{aligned} E[\frac{1}{y+1}] &= \sum_{k=0}^{\infty} \frac{1}{k+1} e^{-\bar{y}} \frac{\bar{y}^k}{k!} \\ &= \frac{1}{\bar{y}} \sum_{k=0}^{\infty} e^{-\bar{y}} \frac{\bar{y}^{k+1}}{(k+1)!} \\ &= \frac{1}{\bar{y}}(1 - e^{-\bar{y}}). \end{aligned} \tag{36}$$

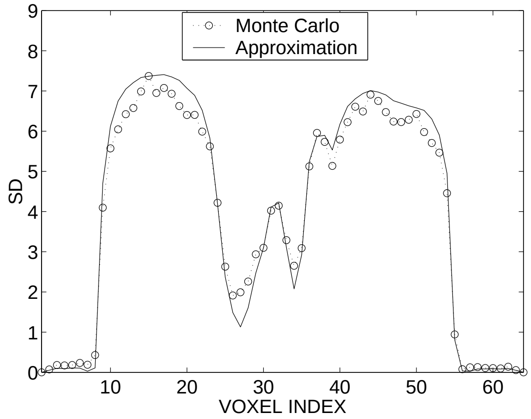


Fig. 14. Transaxial profiles through the CSF region in the second plane of standard deviation images in Fig. 13.

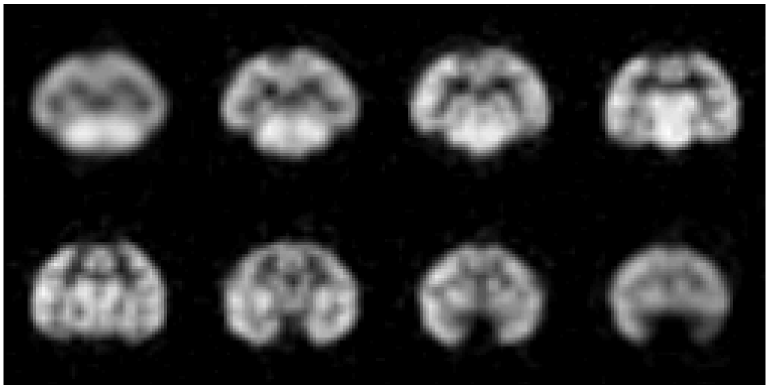


Fig. 15. MAP reconstructions of one baby monkey brain microPET phantom data set with $\beta = 0.0002$.

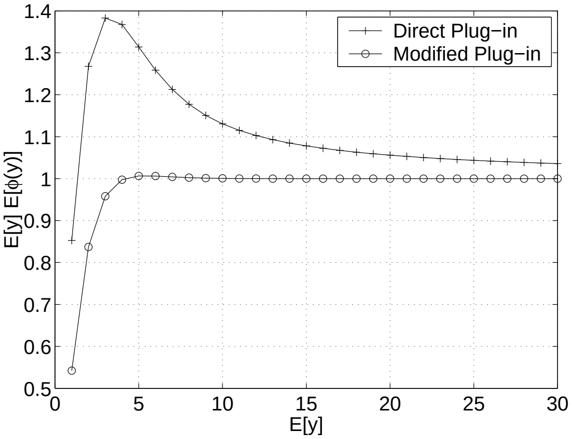
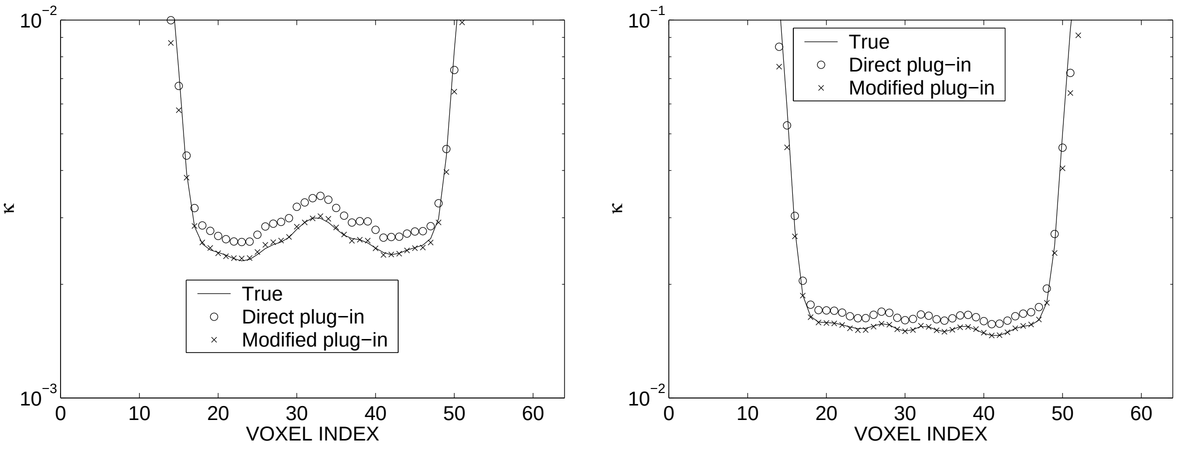


Fig. 16. Plot of bias in the direct and modified plug in methods vs. mean for a Poisson random variable. In each case we plot the product $E[y] \times E[\phi(y)]$ where $\phi(y)$ represents the plug in estimator of $\frac{1}{E[y]}$. Ideally, the product should equal unity.



(a) (b)

Fig. 17. Comparison of the κ values computed using (14) with noise-free projection data (solid line), with the direct plug-in method using noisy data (circle), and with the modified plug-in method using the same noisy data ('x'): (a) the first transaxial plane; (b) the central transaxial plane.

We can therefore use the measurement y_i to compute $\frac{1}{y_i+1}$ as an unbiased estimate of $E[\frac{1}{y_i+1}]$. We then use (36) to compute the corresponding mean $\bar{y}$ via a lookup table. This value is then used to compute κ_j in (14).

The effect on bias of this "modified plug-in" method is shown in Fig. 16. For a Poisson random variable with mean $E(x)$ greater than 3, the method is effectively unbiased estimate. Below this, there is increasing negative bias as the mean value decreases. However the bias is greatly reduced compared to that resulting from the direct plug-in method.

In Fig. 17, we show a comparison of κ_j estimated with this modified plug-in method with the direct plug-in method using the simulated data described in the previous section. We computed the κ_j's using (14) with noise-free projection data, with the direct plug-in method using noisy data, and with the modified plug-in method using the same noisy data. The figure indicates that the modified method corrects most of the bias introduced using the direct plug-in method.

B. Variance Images and ROI quantitation

The voxel-wise sample variances were computed from the 41 reconstructions of the monkey brain phantom and compared to the theoretical approximation results computed using the modified plug-in method. The standard deviation images with selected profiles are shown in Fig. 18. In this case the small Monte Carlo sample size ($L = 41$) results in significantly larger uncertainty

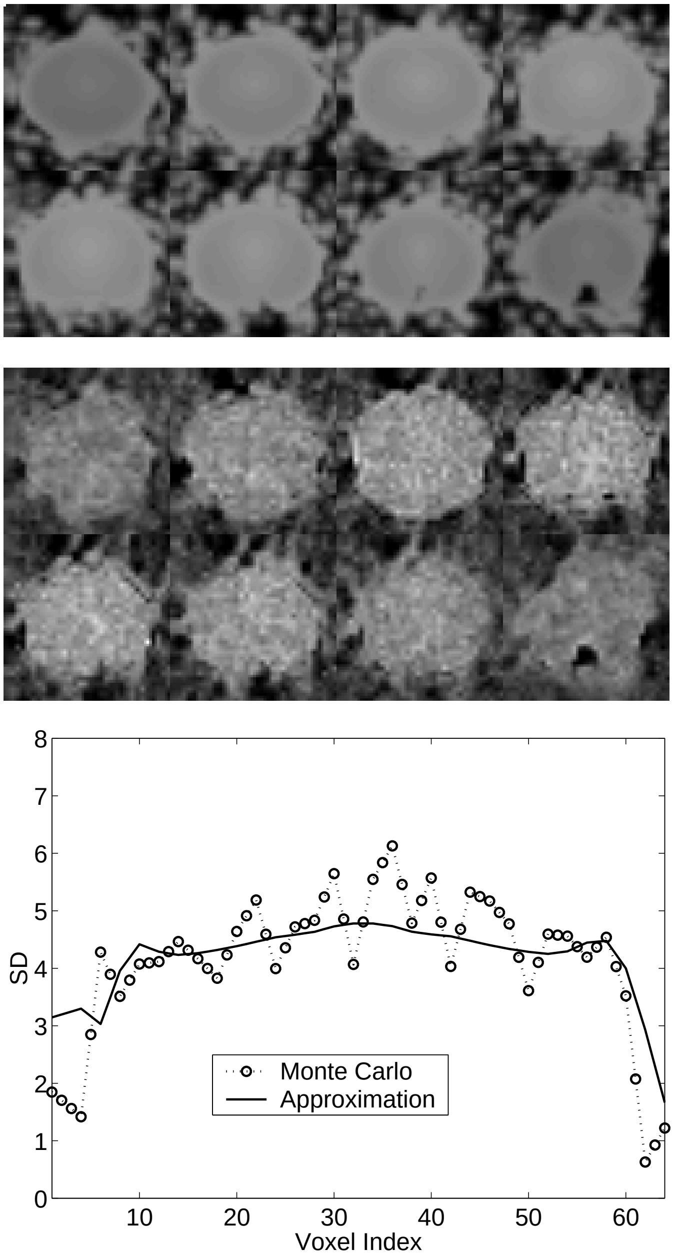


Fig. 18. Standard deviation images of the monkey brain phantom reconstruction. (a) Standard deviation image from theoretical approximation. (b) Standard deviation image from Monte Carlo method. (c) Profiles through the center of the 4th plane.

in the estimated variance than we encountered in the computer simulations with $L = 1000$. To perform a quantitative comparison we looked at activity computed over several ROIs.

We hand selected 21 ROI centers and drew 9 concentric circular ROIs around each selected center by varying the radius from 1 voxel to 9 voxels. As in Section IV-F, we estimated the

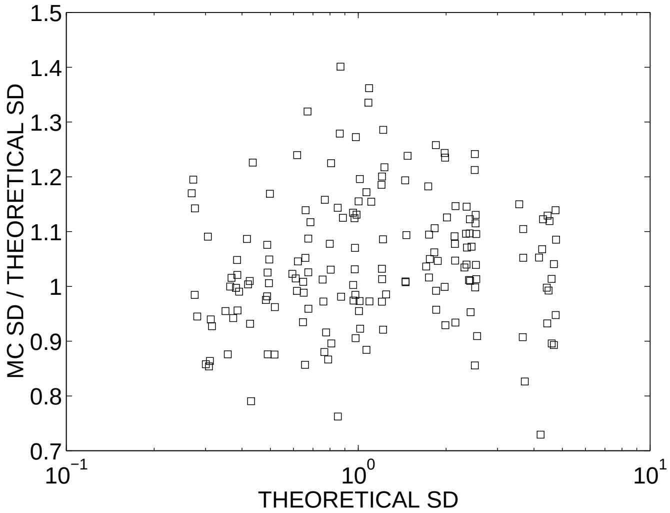


Fig. 19. Ratio of the Monte Carlo estimated standard deviation for each ROI over that computed using the theoretical approximation with the modified plug-in method.

variance of the average activity in each ROI using both theoretical approximation and Monte Carlo methods. The results are shown in Fig. 19. The RMSE between the Monte Carlo result and theoretical approximation in this case is 23.8% with a predicted error of 22.4% error in the Monte Carlo result itself. This result is a practical validation that the variance of ROI quantitation in MAP reconstruction can be estimated to reasonable accuracy in real data when using the modified plug-in method.

VI. CONCLUSION

We have derived simplified expressions for the resolution and noise properties of MAP reconstructions in fully 3D PET. These expressions are rapidly computed and relatively straightforward to interpret. They can be used to characterize the reconstructed images and to optimize system design and reconstruction algorithms. We have also shown how these methods can be used to reconstruct images with near uniform resolution as measured using the contrast recovery coefficient. Extensive Monte Carlo simulations support the accuracy of the approximations used to simplify our theoretical expressions. The experimental phantom scan further confirms these results and demonstrates the use of these methods in plug-in mode.

VII. ACKNOWLEDGMENT

The authors would like to thank Arion Chatziioannou at the Crump Institute for Biological Imaging, UCLA School of Medicine for providing the phantom data shown in Section V, and the reviewers for their thoughtful comments.

REFERENCES

[1] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar, "High resolution 3D bayesian image reconstruction using the microPET small animal scanner," *Physics in Medicine and Biology*, vol. 43, no. 4, pp. 1001–1013, 1998.

[2] J. Qi, R. M. Leahy, C. Hsu, T. H. Farquhar, and S. R. Cherry, "Fully 3D bayesian image reconstruction for ECAT EXACT HR+," *IEEE Transactions on Nuclear Science*, vol. 45, no. 3, pp. 1096–1103, 1998.

[3] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction: Spatial-invariant tomographs," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 1346–1358, September 1996.

[4] P. Bonetto, J. Qi, and R. M. Leahy, "Covariance approximation for fast and accurate computation of channelized Hotelling observer statistics," in *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference*, Seattle, WA, 1999.

[5] H. H. Barrett, D. W. Wilson, and B. M. W. Tsui, "Noise properties of the EM algorithm: I. theory," *Physics in Medicine and Biology*, vol. 39, pp. 833–846, 1994.

[6] W. Wang and G. Gindi, "Noise analysis of MAP-EM algorithms for emission tomography," *Physics in Medicine and Biology*, vol. 42, pp. 2215–2232, 1997.

[7] J. Fessler, "Mean and variance of implicitely defined biased estimators (such as penalized maximum likelihood): Applications to tomography," *IEEE Transactions on Image Processing*, vol. 5, no. 3, pp. 493–506, March 1996.

[8] J. A. Fessler, "Approximate variance images for penalized-likelihood image reconstruction," in *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference*, Albuquerque, NM, 1997.

[9] J. Qi and R. M. Leahy, "A theoretical study of the contrast recovery and variance of map reconstructions from PET data," *IEEE Transactions on Medical Imaging*, vol. 18, no. 4, pp. 293–305, 1999.

[10] J. Qi and R. M. Leahy, "Fast computation of the covariance of map reconstructions of pet images," in *Proceedings of SPIE*, 1999, vol. 3661, pp. 344–355.

[11] M. Yavuz and J. A. Fessler, "Statistical image reconstruction methods for randoms-precorrected pet scans," *Medical Image Analysis*, vol. 2, no. 4, pp. 369–378, 1998.

[12] E. Mumcuoglu, R. Leahy, S. Cherry, and E. Hoffman, "Accurate geometric and physical response modeling for statistical image reconstruction in high resolution PET," in *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference*, Anaheim, CA, 1996, pp. 1569–1573.

[13] J. A. Fessler, "Preconditioning methods for shift-variant image reconstruction," in *Proc. IEEE International Conference on Image Processing*, Santa Barbara, CA, 1997, vol. 1, pp. 185–188.

[14] J. A. Fessler and S. D. Booth, "Conjugate-gradient preconditioning methods for shift-variant image reconstruction," *IEEE Transactions on Image Processing*, to appear.

[15] S. R. Cherry, Y. Shao, S. Siegel, R. W. Silverman, K. Meadors, J. Young, W. F. Jones, D. Newport, C. Moyers, E. Mumcuoglu, M. Andreaco, M. Paulus, D. Binkley, R. Nutt, and M. E. Phelps, "MicroPET: A high resolution PET scanner for imaging small animals," *IEEE Transactions on Nuclear Science*, vol. 44, no. 3, pp. 1161–1166, June 1997.

[16] J. A. Fessler, "Penalized weighted least squares image reconstruction for PET," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 290–300, 1994.

[17] C. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Transactions on Image Processing*, vol. 5, no. 3, pp. 480–492, March 1996.

[18] E. Mumcuoglu, R. Leahy, S. Cherry, and Z. Zhou, "Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 687–701, December 1994.

[19] R. E. Carson, Y. Yan, M. E. Daube-Witherspoon, N. Freedman, S. L. Bacharach, and P. Herscovitch, "An approximation formula for the variance of PET region-of-interest values," *IEEE Transactions on Medical Imaging*, vol. 12, no. 2, pp. 240–250, 1993.

[20] E. Hoffman, P. Cutler, W. Digby, and J. Mazziotta, "3D phantom to simulate cerebral blood flow and metabolic images for PET," *IEEE Transactions on Nuclear Science*, vol. 37, pp. 616–620, 1990.

[21] D. W. Wilson, B. M. W. Tsui, and H. H. Barrett, "Noise properties of the EM algorithm: II. Monte Carlo simulations," *Physics in Medicine and Biology*, vol. 39, pp. 847–872, 1994.

[22] B. R. Frieden, *Probability, Statistical Optics, and Data Testing*, Springer-Verlag Berlin Heidelberg New York, 1983.